

Benchmarks for solar radiation time series forecasting

Cyril Voyant^{a,*}, Gilles Notton^a, Jean-Laurent Duchaud^a, Luis Antonio García Gutiérrez^a,
Jamie M. Bright^b, Dazhi Yang^c

^a SPE Laboratory - UMR CNRS 6134 - University of Corsica Pasquale Paoli - Route des Sanguinaires, F20000 Ajaccio, France

^b UK Power Networks, London, UK

^c Harbin Institute of Technology, School of Electrical Engineering and Automation, Harbin, Heilongjiang, China

ARTICLE INFO

Article history:

Received 3 January 2022

Received in revised form

14 March 2022

Accepted 10 April 2022

Available online 18 April 2022

Keywords:

Irradiation

Filtering

Exponential smoothing

Combination

Benchmark

Forecast

ABSTRACT

With an ever-increasing share of intermittent renewable energy in the world's energy mix, there is an increasing need for advanced solar power forecasting models to optimize the operation and control of solar power plants. In order to justify the need for more elaborate forecast modeling, one must compare the performance of advanced models with naïve reference methods. On this point, a rigorous formalism using statistical tools, variational calculation and quantification of noise in the measurement is studied and five naïve reference forecasting methods are considered, among which there is a newly proposed approach called ARTU (a particular autoregressive model of order two). These methods do not require any training phase nor demand any (or almost no) historical data. Additionally, motivated by the well-known benefits of ensemble forecasting, a combination of these models is considered, and then validated using data from multiple sites with diverse climatological characteristics, based on various error metrics, among which some are rarely used in the field of solar energy. The most appropriate benchmarking method depends on the salient features of the variable being forecast (e.g., seasonality, cyclicity, or conditional heteroscedasticity) as well as the forecast horizon. Hence, to ensure a fair benchmarking, forecasters should endeavor to discover the most appropriate naïve reference method for their setup by testing all available options. Among the methods proposed in this paper, the combination and ARTU statistically offer the best results for the proposed study conditions.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

Leveraging physics-based and statistical methods to explain and forecast time series has been playing an increasingly important role in energy meteorology [1,2]. It is logical to expect a rapid development of forecasting methods, in terms of both number and intricacy [3]. In order to justify the choice of one method over another, forecast comparison is necessary. In that, superiority claims ought to be regarded as an essential component in all forecasting works. Under the “publish or perish” regime, overarching statements such as “the proposed method outperforms the existing ones” are ubiquitous and rapidly updated, which tends to result in a literature that is difficult to follow, to interpret, and to condense. Precautions must be taken.

According to Wolpert's “no free lunch” theorem [4], each time a

predictive method is deemed superior in some aspects, there ought to be some other aspects in which the method does not perform as well comparing to its alternatives. In this regard, one may consider the classic example given by Gneiting [5], in which two sets of forecasts optimized under different objectives (least squares and least absolute deviations) outperform each other under different evaluation metrics (root mean square error and mean absolute error); this is formally known as *consistency in forecast verification* [6]. In a more general sense, even when a simple method is compared to a sophisticated method, the former method has the advantage of being easy to implement, although its accuracy may not be as attractive as the latter. On this point, the problem of choosing a forecasting method, like many other problems of a similar nature, is one of balance.

Another important aspect of forecast comparison is the situation under which the verification is conducted. In the field of renewable energy, in particular, intermittent solar energy, the optimality of a method depends on the geographical location, climate and weather regime [7], time-step, and forecast horizon; it does not make much

* Corresponding author.

E-mail address: cyrilvoyant@gmail.com (C. Voyant).

sense to compare directly two sets of forecasts issued under distinct forecasting situations. Suppose some forecasts have been reported to be optimal in one forecasting situation, it may not compare favorably against its peers in another situation [8]. Stated differently, knowing the optimal choice of method to be used in a particular situation is not straightforward, and one cannot simply take as true the conclusions made elsewhere. This is a limitation as to the general validity of inductive reasoning. In that, the only logical way to confirm that a method is indeed optimal is to test all existing methods for that given forecast situation.

This task is a necessarily difficult one. It is for that reason that, until now, there is no consensus on what constitutes a perfect model [9–11]. Since enumerating and testing all existing methods is never quite possible, it is customary to leverage a naïve reference method during forecasting, such that forecasts made at different locations and over different time periods can be compared, though with appropriate caveats. Depending on the amount of improvement acquired, there would be different levels of enthusiasm for the methods of interest. In order for this approach to take effect, most generally, several prerequisites must be evaluated carefully. Firstly, reference methods should be applicable in various independent and operational modes [12–15]. That is to say that the reference method should be “universal” in a sense that it does not depend on the type of available data. Secondly, the reference method must be sufficiently naïve, in that, it does not require dedicated knowledge to implement. For this reason, simple or very simple reference methods are preferred over elaborate reference methods that only provide slightly better results. Thirdly, when multiple naïve reference methods are present, the one with the highest accuracy should be used [16].

There are many reference methods in weather forecasting [17,18]. The most widely used ones are climatology and persistence [19] though the accuracy of these two references are generally low. Other time series reference methods are occasionally used, such as the simple exponential smoothing [20]. If we were to expand the list of reference methods in that direction, the choice becomes more flexible. In this paper, we propose to explore further options in order to effectively judge the quality of any new or existing forecasting methods using time series formalism in the solar radiation field. A new baseline approach is introduced to have a global indicator of the performance of all the methods. Among these methods, there are regressive methods, stochastic learning methods, deep learning methods, genetic algorithm, data-driven methods, local-sensing methods, hybrid forecasting methods, and application orientated methods.

The following pages are organized as follows. In Section 2, various reference forecasting methods are introduced, which include four well-known methods, a newly proposed one, and a linear combination of the previous five. In Section 3, the specifics related to the forecasting of climate and weather time series are underlined. In Section 4, an application of these methods is firstly demonstrated on global horizontal irradiance (GHI) data measured in Ajaccio, France. Then, we check if the conclusions obtained for this site can be generalized for GHI series from four other sites with different climatological characteristics. The forecasting methods are applied to tilted global irradiance (TGI), which is more useful for solar photovoltaic applications, in view to verify the relevance of the methods. The final results refer to the temperature and wind time series in order to judge the portability of the methods we want to qualify as reference. Finally, Section 5 is dedicated to conclusions.

As is discussed further in the next section, it is crucial to develop the notion of statistical reference method (SRM) for solar radiation forecasting. It is in this context, that we revisit classical methods that are either widely or not widely used resulting in two new proposed methods based on a novel formalism.

2. Statistical reference methods (SRM)

The existing methodologies that are most used in the context of solar resources forecasting, to facilitate integration and management of photovoltaic (PV) plants, smart grid control, or energy trading on the spot market, are based on the time series formalism and numerical weather prediction (NWP). Comparing these two classes of methods, the former is based on extrapolation of data, whereas the latter solves the governing partial differential equations which describes the state of the atmosphere. Since the time series approaches are far less computational demanding than NWP, they are more attractive in terms of forecast benchmarking. In this paper, all benchmarking methods considered are in accordance with time series approaches; they are abbreviated as statistical reference methods (SRMs).

Considering a signal x with samples x_t being regularly-spaced values in time, and transforming it such that $x_t^c \equiv x_t - \mathbb{E}[x]$, which implies the transformed variable has expected value: $\mathbb{E}[x^c] = 0$. One fundamental way to describe such a signal is to assume the forecast is a function of the most recent observation(s), or an autoregressive (AR) model. For a forecast at horizon h and a direct multi-step forecast strategy,¹ it is expressed:

$$x_{t+h}^c = \alpha x_t^c + \omega_{t+h}, \quad (1)$$

where x_t^c is a centered time-dependant signal, α ($|\alpha| < 1$ for process stability) is a gain term and ω is additive noise. The direct strategy could be applied because in solar forecasting context this is often the best method although there is no real consensus on this subject. As mentioned by Taieb and Hyndman [22], “Choosing between these different strategies involves a trade-off between bias and estimation variance”. Under the state space framework in the discrete domain where Kalman filter is often used, x_t^c could be defined as the state vector of the process at time t , and α the state transition parameter of the process from the state at t to the state at $t + h$, assuming stationarity over time. ω is the associated zero-mean white noise process, i.e. $\omega \sim \mathcal{N}(0, \sigma_\omega^2)$ with $\sigma_\omega^2 = \mathbb{E}[\omega_t^2]$.

With this in mind, observations on this variable are contained in the series y_t . The centered form, denoted y_t^c can be modeled by the observation equation:

$$y_t^c \equiv y_t - \mathbb{E}(y) = x_t^c + v_t, \quad (2)$$

where y_t is the actual measurement of x at time t and v_t is the associated measurement error with $v \sim \mathcal{N}(0, \sigma_v^2)$. Note that the hypothesis $\mathbb{E}(x) = \mathbb{E}(y) \equiv \bar{y}$ is implied in Eq. (2) and therefore, a bias in the measurement would be detrimental: a quality control of the sensors is a prohibitive prerequisite.

2.1. Usual benchmark methods for meteorological time series

There are many reference methods in meteorology and mainly in solar radiation. In this section and in Appendix D, we will list the most commonly used ones.

2.1.1. Persistence (PER)

Persistence is the simplest case where $\alpha = 1$ in Eq. (1), the prediction of x_{t+h}^c denoted \hat{x}_{t+h}^c becomes:

¹ to contrast with recursive multi-step forecast [21] where the prediction for the prior time step is used as an input for making a prediction on the following time step.

$$\hat{x}_{t+h}^c = x_t^c \quad (3)$$

which implies that $\hat{x}_{t+h} = y_t$. This model (see [Appendix D.1](#) for details) is very reliable when the meteorological data series has a very low variability but it quickly becomes ineffective with very noisy, periodic or trending signals. However, in operational cases (forecasting for energy or power management), when the user needs a forecast and when there is no measurement history, it is often the only way to proceed in the nowcasting case (from 1h to 6h). For benchmarking at larger horizons, a solution (and undoubtedly the best alternative) would be the use of climatology.

2.1.2. Climatology (CLIM)

If in the previous case the method used was only valid for strongly persistent phenomena (low variability), the model presented here corresponds to the case where there is no statistical dependence between the different measurements (i.e. white noise). In this case, we should choose $\alpha = 0$ where it immediately becomes $\hat{x}_{t+h}^c = 0$ and therefore $\hat{x}_{t+h} = \mathbb{E}(y)$ (historical mean). This model (see [Appendix D.2](#) for details), though simple, should not be overlooked, because as soon as the autocorrelation coefficient ρ comes close to 0, it is often the best way to attain a minimum forecast error. This model becomes interesting when predicting meteorological time series, and when deep horizons close to the predictability limit of chaotic phenomena are studied.

2.1.3. Autoregressive model of order one AR(1) or climatology persistence (CLIPER)

In this section the link between the AR(1) estimate and the well-named “climatology–persistence combination,” or CLIPER for short, is shown. This kind of model can be considered as a reference because it is easy to implement and does not require any learning phase. Hence, it is easy to use and robust in case of data are suddenly missing due to a detector failure if the forecasts are used automatically in operational mode. Now, consider the Yule–Walker equations [23] or equivalently multiply Eq. (1) by x_t^c and take the expectation value. x_t^c is considered as a stationary process in the weak sense in the rest of the paper (also denoted wide-sense stationarity, or covariance stationarity). It follows that:

$$\mathbb{E}[x_{t+h}^c x_t^c] = \mathbb{E}[\alpha x_t^c x_t^c + \omega_{t+h} x_t^c]. \quad (4)$$

After some mathematical considerations detailed in [Appendix A](#) and using the correlation factor $\rho(h)$, it comes that the prediction can be expressed by:

$$\hat{x}_{t+h} = \rho(h)y_t + [1 - \rho(h)]\mathbb{E}(y). \quad (5)$$

An alternative, and shorter, way to arrive at Eq. (5) can be found in Ref. [14], in which the same results can be obtained by minimizing the mean square error (MSE) of CLIPER forecasts. The pseudo-code used to predict with CLIPER is given in [Appendix D.3](#).

2.1.4. Simple exponential smoothing (ES)

Forecasts are calculated using weights that decrease exponentially as data become older (where $|\alpha| < 1$ is the smoothing parameter) as described in Eq. (6) [20].

$$\hat{x}_{t+h}^c = \alpha x_t^c + \alpha(1 - \alpha)x_{t-1}^c + \alpha(1 - \alpha)^2 x_{t-2}^c + \dots \quad (6)$$

That is, all forecasts take the same value (h -step-ahead forecast is constant $\forall h$), equal to the last level component. Remember that these forecasts are only suitable if the time series has no trend or seasonal component. Applying the same tools exposed in the beginning of Section 2, Eq. (6) can be replaced with:

$$\hat{x}_{t+h} - \bar{y} = \alpha(y_t - \bar{y}) + \alpha(1 - \alpha)(y_{t-1} - \bar{y}) + \alpha(1 - \alpha)^2(y_{t-2} - \bar{y}) + \dots \quad (7)$$

The separation of the \bar{y} and y_{t-i} terms leads to:

$$\begin{aligned} \hat{x}_{t+h} = & \bar{y}[1 - \alpha - \alpha(1 - \alpha) - \alpha(1 - \alpha)^2 - \dots] + \alpha y_t + \alpha(1 - \alpha)y_{t-1} \\ & + \alpha(1 - \alpha)^2 y_{t-2} + \dots \end{aligned} \quad (8)$$

If the second part of this equation (Eq. (8)) can be reduced according to the sum $\sum_{i=0}^n \alpha(1 - \alpha)^i y_{t-i}$, the first term is related to the sum of the first $n + 1$ terms of a geometric series (common ratio = $1 - \alpha$). In the end, the ES model is given by:

$$\hat{x}_{t+h} = \alpha \sum_{i=0}^n (1 - \alpha)^i y_{t-i} + \bar{y}(1 - \alpha)^{n+1}, \alpha \neq 0. \quad (9)$$

In the exponential smoothing method (see [Appendix D.4](#) for details), the predictions lie between the two extremes, the naive persistence $\hat{x}_{t+h} = y_t$ (when $\alpha = 1$) and the simple average $\hat{x}_{t+h} = \bar{y}$ (when $\alpha = 0$), which assumes that all observations are of equal importance, and assigns them equal weights when generating forecasts. In order to be consistent with other methods, we choose to determine α without using an optimization phase. As such, considering that $\alpha = \rho(1)$ we obtain a predictor close to the persistence when $\rho(1)$ is close to 1 and close to the mean when it tends to 0. Note that we could use the relation $\alpha = \rho(h)$ and this could make sense but we wanted to present the simplest method relating to exponential smoothing.

2.2. Proposed methodologies

The methods used as reference or so-called naive methods in terms of weather forecasting through the time series formalism are often (and this is the entirely the purpose) very easy to implement. Qualifying the following methods as naive may seem counterintuitive given the calculations required to develop it. However, they are only ever performed once, then methods are applicable very simply to whatever the studied time series (solar radiation or not).

2.2.1. Particular autoregressive model of order two (ARTU)

In this part, we propose a methodology that improves the performances of the AR(1) model (or CLIPER) previously discussed in Section (2.1.3). We show that, even if this is not part of the initial hypotheses [see Eq.10 and 11], this approach is equivalent to an AR(2), with the difference being that the estimation of the coefficients is only based on the autocorrelations generation—since the method is a particular form of AR(2), it is called ARTU, which is pronounced as “A-R-two.” Firstly, it is useful to think of another version of Eq. (1), in which the state estimate is known to be *not* optimal (x_{t+h}^{c-}).

$$x_{t+h}^{c-} = \alpha x_t^c + \omega_{t+h}. \quad (10)$$

We further assume the need for an update (\hat{x}_{t+h}^{c-}) related to a linear combination between \hat{x}_{t+h}^{c-} ($= \alpha x_t^c$) and the previous innovation (or residual i.e. $y_t - \hat{x}_t^{c-}$). Shown in Eq. (11) is the updated form, whereby the factor K can be treated as a gain, and the combination as a filtering process.

$$\hat{x}_{t+h}^c = \hat{x}_{t+h}^{c-} + K(y_t - \hat{x}_t^{c-}). \quad (11)$$

Note that this approach, although it is quite close to what is observed with prediction of a classical ARMA(1,1) model [24] or a Kalman filtering [25], is different. In the first case, ω_{t+h} would be the residual of \hat{x}_t^c and not \hat{x}_t^{c-} , moreover K would be multiplied by

($y_t^c - \hat{x}_t^c$) in Eq. (11). The second one would require a modification in Eq. (11), in that, K would be multiplied by $(y_{t+h}^c - \hat{x}_{t+h}^c)$. The assumptions of stationarity formulated in Section 2.1.3 are maintained. From Eq. (10) and Eq. (11), the forecast can therefore be put in the following form:

$$\hat{x}_{t+h}^c = \alpha x_t^c + K(y_t^c - \alpha x_{t-h}^c). \quad (12)$$

The transition between states and measurements is:

$$\hat{x}_{t+h} - \bar{y} = \alpha(y_t - \bar{y}) + K[y_t - \bar{y} - \alpha(y_{t-h} - \bar{y})], \quad (13)$$

$$\hat{x}_{t+h} = (K + \alpha)y_t - (K\alpha)y_{t-h} + (1 + K\alpha - K - \alpha)\bar{y}. \quad (14)$$

A more practical parameterization can be proposed by letting $P = K\alpha$ and $S = K + \alpha$. This allows us to present \hat{x}_{t+h} as a convex combination of y_t , y_{t-h} and $\mathbb{E}(y)$ from the sum S and the product P :

$$\hat{x}_{t+h} = Sy_t - Py_{t-h} + (1 + P - S)\bar{y}. \quad (15)$$

Using what has just been presented above and in Appendix B, the optimal values of α and K can be determined by minimizing the MSE of ARTU forecasts. Based on the results therein, the optimal α and K can be written in functions of a triplet $[R, \rho(h), \rho(2h)]$, where $\rho(h)$ and $\rho(2h)$ are two correlation coefficients that have already defined previously, and $R = \sigma_v^2 / \sigma_x^2$ is linked to the quality of the measurement. The ARTU prediction is described by the pseudocode in Appendix D.5. There are, nevertheless, some hypotheses that have been formulated, especially concerning the stationarity of the studied time series. Indeed, wide-sense stationary assumption (mean and variance are time-independent and autocovariance and autocovariance can be expressed as functions of the time-lag) is necessary to obtain a closed-form in B.6 allowing to simply propose S and P in Eq. (15). This is important to keep in mind, because if we deviate too much from these conditions (that can be qualified as ideal), there is a good chance that the ARTU approach (prediction with Eq. (14) according the α and K parameters obtained solving Eq.(B.6)) does not give satisfactory results. To arrive at an approximately stationary time series in respect to solar radiation, it is customary to use a clear-sky model [26–28], which is often simplified from the Beer–Lambert relationship (See Section Appendix C). We would see that the non-stationary phenomenon appears during the temperature and wind speed studies (in Section 4.4) where there is no simple knowledge model, where using a ratio to trend allows us to get closer to the ideal and stationary case.

We computed α and K for values of $\rho(h)$ and $\rho(2h)$ from -1 to $+1$ (with an incremental step of 0.01) and for 4 values of R ($0, 0.01, 0.05$ and 0.1) solving Eq.(B.6). An example is given in Fig. 1 and other values are available in <https://github.com/cyrilvoyant/ARTU.git> (via Matlab® codes). What is important to remember is that once these values are known, the forecasts become relatively simple to implement by knowing the autocorrelations of the series studied. The forecast is direct, fast and does not require modeling or learning. The parameters are obtained by data mining approach and the method can then be qualified as naïve although a minimum of historical data is necessary.

A potential limitation on the use of the ARTU approach can be enacted from Eq. (11). Indeed, by multiplying by x_{t-h}^c and taking the expectation value, it comes Eq. (16) where we observe a condition between the two correlations of the model.

$$\rho(h)^2 - \rho(2h) = 0. \quad (16)$$

This means that predictions can be efficient for autocorrelation factors (ACF) that satisfy this equation but becomes less reliable as soon as we deviate from this case. Note that $\rho(h) = a^{bh}$ is a solution

of this equation ($a, b \in \mathbb{R}$ with $\log(a)b < 1$) with the constrain of $|\rho(h)| < 1$.

As soon as the ACF deviates from the exponential decay, the results are best with the filtering proposed here while when the decay is respected the filtration is not necessary. This induces $K = 0$ and $\alpha = \rho(h)$ and the model becomes equivalent to the classical CLIPER presented in Section 2.1.3. In summary, the best results are observed without filtration and in the presence of the exponential decay of the ACF, but when this one is not observed, the filtration takes all importance and in theory improves the predictions.

Before going any further, it is thought essential to clarify and discuss the variable R . From its construction, we realize that the variance ratio fluctuates between 0 ($\sigma_x^2 \gg \sigma_v^2$) and 1 ($\sigma_x^2 \sim \sigma_v^2$). Usually, the measuring devices used in meteorology are quite efficient, so we can limit the values of R between 0 and 10% of σ_x^2 . According to Vuilleumier et al. [29], one can logically imagine using an R between 1% and 5% .

Fig. 1 shows as example (for $R = 0.05$) the values of α and K obtained with the methodology exposed previously. Four areas are visible concerning K , bounded by the line $\rho(h) = 0$ and the parabola defined by the equation $\rho(2h) = \rho(h)^2$. Magnification of the area related to positive value of $\rho(h)$ and $\rho(2h)$ is proposed in Fig. 2. Positive values correspond to what is usually observed in solar radiation. All the files used to generate the α and K coefficients are available in <https://github.com/cyrilvoyant/ARTU.git> ($\mathcal{M}(R)$ matrices and codes). For a forecast at the horizon h , the procedure is as follows:

- ↔ Calculate $\rho(h)$ and $\rho(2h)$ using in-sampling data;
- ↔ Obtain the α and of K values by consulting on Figs. 1 and 2 or generating them more precisely using the code available in Appendix D.5 and the matrices ($\mathcal{M}(R)$);
- ↔ Derive forecast using Eq. (15).

In the end, let us suppose that we seek to use the filtration previously stated, for a signal concerning $h = 1$, if one estimates $\rho(1) = 0.4$ and $\rho(2) = 0.3$; we therefore have $\hat{x}_{t+1} = Sy_t - Py_{t-1} + (1 + P - S)\bar{y}$. Cf. Eqs. 11–15, and so:

- ↔ $\hat{x}_{t+1} = 0.33y_t + 0.16y_{t-1} + 0.51\bar{y}$ for $R = 0.01$, $\alpha = 0.60$ and $K = -0.27$ hence $S = 0.33$ and $P = -0.16$;
- ↔ $\hat{x}_{t+1} = 0.34y_t + 0.15y_{t-1} + 0.51\bar{y}$ for $R = 0.05$, $\alpha = 0.59$ and $K = -0.25$ hence $S = 0.34$ and $P = -0.15$;
- ↔ $\hat{x}_{t+1} = 0.35y_t + 0.13y_{t-1} + 0.52\bar{y}$ for $R = 0.10$, $\alpha = 0.58$ and $K = -0.23$ hence $S = 0.35$ and $P = -0.13$.

2.2.2. Combination of methods (COMB)

According to the conclusions related to the M4 forecasting competition [30], one of the major findings is related to the combination of methods [31]. In fact, among the most accurate methods, the vast majority (12 of the 17 most accurate) were a combination of statistical approaches. Moreover, this aspect is so important that one innovation was the introduction of a combination reference for benchmarking the accuracy of the methods submitted to the competition. It therefore seems logical to draw inspiration from this remarkable conclusion in order to propose a reference predictive methodology based on the combination of the simple methods presented in this study. Still inspired by Ref. [30] we chose to use the simple arithmetic average of the different outputs of the models although the use of the median may show equally good results in some situations [32]. Note that the term “combination of methods” refers to the statistician’s universe while machine learning researchers call it “ensemble learning” or “classifier combination”.

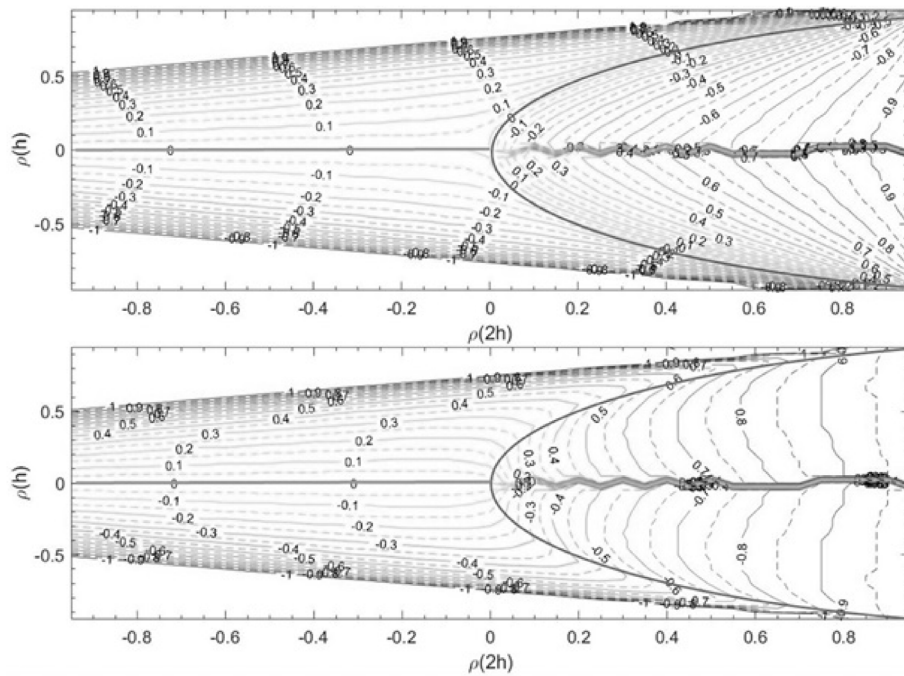


Fig. 1. K (top) and α (bottom) values according to the values of $\rho(h)$ and $\rho(2h)$ and for $R = 0.05$ (resolution of Eq.(B.6)).

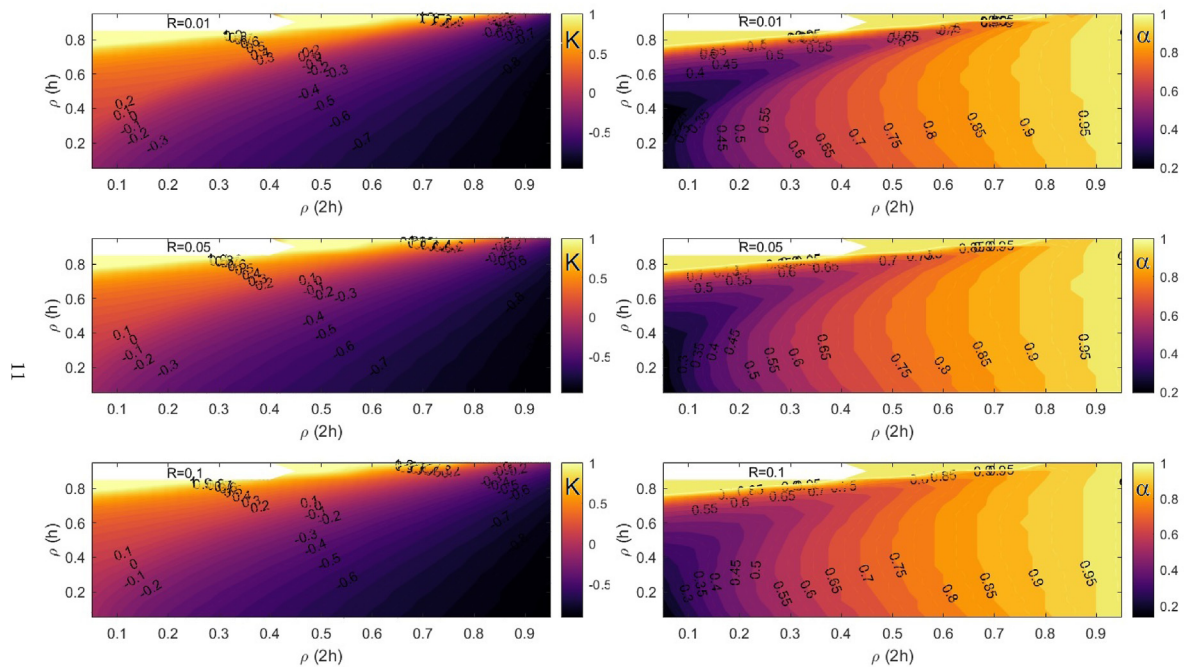


Fig. 2. α and K values according to the values of $\rho(h)$, $\rho(2h)$ and R .

Whatever the name, this tool is often useful in dealing with potential sources of non-stationary variation present in the data. The Timmermann' paper [33] allows to understand but especially to validate the use of COMB.

3. Experimental setup

To compare the SRMs presented in this paper as objectively as possible, we must establish rules follow best practices in the fields of meteorology, renewable energy, and particularly, solar irradiance

forecasting. The seasonality of considered time series and ad-hoc test (based on auto-correlation) are discussed in Appendix C. In what follows, I_{GH} and I_{CS} are used to denote GHI and its clear-sky expectation, whereas \hat{I}_{GH} is the GHI forecast. To be more precise, the clear-sky index, which is defined as $\kappa = I_{GH}/I_{CS}$, is used as x as in Eq. (1).

3.1. The pre-treatment

As part of this study, several rules and explanations must be given to improve the objectivity of the solar energy forecasting

conclusions:

→ Irradiance time series (I_{GH}) are measured in several sites around the world with different climatic characteristics. For each of them, the Köppen–Geiger classification (KG) [34], the expected forecastability (F) [35] and the geographic coordinates are provided. At least two years of data are available for each site and used during the simulations. Usually one year of data is considered sufficient to validate a forecasting method, though two years are preferable;

→ The models are evaluated during daytime irradiance values only, filtering the checked data (according to quality control [36,37]) of the solar zenith angle (I_{GH} where zenith angle $\theta_Z > 85^\circ$ are excluded);

→ The data was acquired every 15 min or every hour, and the corresponding forecast horizons are between 15 and 150 min, in the first case, and between 1 and 10 h, in the second case.

3.2. Error metrics

In order to compare the accuracy of the forecasting methods, we refer to the official accuracy measures of the M4 competition, i.e., the mean absolute scaled error (MASE [38]) in Eq. (17) computed for seasonal time series in retrospective case. MASE admits an average value as denominator, so as long as the filtering of the night hours is operated (see previous subsection), the denominator in Eq. (17) will be neither equal nor close to 0. This metric should not be used alone, it is only computed to improve and validate results obtained with the usual methods (Eqs. 18 and 19 [6]). Error is calculated on the out-sample data and averaged over all horizons, the main interest of MASE as defined in Eq. (17) is tied with that one coefficient for all horizons (retrospective case with multiple step ahead forecasts).

$$\text{MASE} = \frac{100}{h} \frac{\sum_{t \in \text{Test}} \sum_{i=1}^h |I_{GH}(t+i) - \hat{I}_{GH}(t+i)|}{\sum_{t \in \text{Test}} |I_{GH}(t) - I_{GH}(t-m)|}, \quad (17)$$

where m is the period, see Eq. (C.2), and Test indicates the test sample of size n with $n \gg m$. The MASE method requires a normalization, performed here using the signal period. The filtration presented previously (Section 3.1) has an adverse effect where it reduces the number of exploitable data but also modifies the seasonality (not constant during the year). As a result, it is possible that the denominator of Eq. (17) takes quite high values, which may result in quite low MASE. As all methods are evaluated identically, the interpretation of MASE remains valid, besides a non-periodic version could have been used.

If these metrics are the most commonly used by researchers working on time series forecasting, the analysis of the literature shows that they are only rarely (if ever) used to validate predictions related to global radiation or photo-voltaic power. There are simpler ones that are used in meteorology and more particularly in deterministic solar resource forecasting [6], i.e., the normalized mean absolute error (nMAE), see Eq. (18) and the normalized root mean square error (nRMSE), see Eq. (19) [39].

$$\text{nMAE}(h) = 100 \frac{\sum_{t \in \text{Test}} |I_{GH}(t+h) - \hat{I}_{GH}(t+h)|}{\sum_{t \in \text{Test}} I_{GH}(t)}, \quad (18)$$

$$\text{nRMSE}(h) = 100 \sqrt{n} \frac{\sqrt{\sum_{t \in \text{Test}} [I_{GH}(t+h) - \hat{I}_{GH}(t+h)]^2}}{\sum_{t \in \text{Test}} I_{GH}(t)}. \quad (19)$$

Keeping in mind that ARTU was built from the minimization of

the L^2 norm error function (MSE), it would seem logical in the following that the contribution of the filtering is more consistent with nRMSE than with nMAE but it is imperative to quantify what is happening for both.

4. Results

All simulations were performed with using Matlab and were run on a standard personal computer (Intel core i7, 16 GB RAM). Reports of the execution times are omitted throughout, since the times are very short, due to the fact that the proposed models are “naïve,” i.e., there is no learning phase.

4.1. Specific location

In this part, only the results obtained in France at the Ajaccio site (west coast of Corsica 41°55'36"N, 8°44'13"E, alt 30 m) for different forecast horizons are shown. The climate is Mediterranean, with mild, relatively rainy winters and hot, sunny summers, sometimes sultry, but tempered by the breeze ($KG = \text{Csa}$). The site is in an area exposed to the Mistral wind, which blows from the Gulf of Lion. The average temperature of the coldest months (January, February) is 9 °C, that of the warmest months (July, August) 22.5 °C. The forecastability F is estimated close to 68% [35], which gives it climatic characteristics relative to cloudy occurrences (and thus to solar radiation) relatively straightforward to predict.

It is important to verify the impact of the ratio to seasonal trend. According to Eq.(C.2) and considering a 90% confidence level ($\alpha = 0.1$), the quantile $q_{0.95} = 1.645$. Computing ρ with 100 data points ($n = 100$) we obtain for global solar irradiation (I_m) $t(m) = 0.4032$ and $\rho(m) = 0.62$ while in the clear sky index case (κ_m), $t(m) = 0.2918$ and $\rho(m) = 0.1842$. The use of clear-sky series has a significant impact on seasonality, and we therefore consider that the predictive methodology described above can be applied, as soon as the ratio to trend (I_{CS}) is carried out.

Tables 1 and 2 show the error metrics related to prediction concerning horizon between 1 and 10 h. The presented models are persistence (PER in Section 2.1.1), climatology (CLIM in Section 2.1.2), climatology persistence (CLIPER or AR(1) in Section 2.1.3), exponential smoothing (ES in Section 2.1.4), our proposed modified ARTU (or AR(2) in Section 2.2.1) and a combination of CLIPER, ARTU (for $R = 0.05$) PER and ES (COMB in Section 2.2.2). The MASE is computed with $m = 13$ and not 24 because the filtration of night hours reduces the number of data per day, depending on the day considered (time of year); the periodicity varies between 9 and 15 h on this site.

These tables provide a lot of information that should be confirmed by simulation of the models at other experimental sites. It is important to remember that the goal of the predictive methodologies described and tested here is not to be the best forecasting models, but the simplest ones (sometimes very simple) that would allow arbitration of the classification of more complex models. For example, forecasts made with artificial neural networks of the multilayer perceptron type on this data set are slightly less than 20% for 1-h horizon [40] but are too complex to be used as reference.

First of all, climatology is not a good predictive model, but it is the simplest to implement (nRMSE > 70% whatever the horizon studied). At the Ajaccio site, persistence is a very good indicator for short horizons, but loses its predictive power from $h > 4$. Since many energy applications (such as energy management systems of smart or micro grids) focus on horizons lower than 6 h, this model is not excluded. If we want to improve the results, climatology-persistence is a very good alternative, followed closely by exponential smoothing. The ARTU model gives systematically better results than the climatology-persistence as soon as $R > 0$ but in the specific case of the studied site the gain is minimal. In the following

Table 1
nRMSE and nMAE for the six different benchmark methods at Ajaccio, France. The lowest error metrics values for each horizon are highlighted in bold.

		Horizons (h)									
		1	2	3	4	5	6	7	8	9	10
PER	nRMSE	22.1	31.8	40.2	47.6	52.7	55.0	55.2	53.6	51.7	49.3
	nMAE	11.9	17.6	21.7	25.1	27.6	29.3	30.1	30.3	29.9	29.2
CLIM	nRMSE	70.6	70.6	70.7	70.7	70.7	70.8	70.8	70.8	70.8	70.8
	nMAE	60.7	60.8	60.8	60.8	60.9	60.9	60.9	61.0	61.0	61.0
CLIPER	nRMSE	21.3	27.7	30.5	32.5	33.6	34.0	34.3	34.5	34.7	34.8
	nMAE	13.9	18.8	20.7	21.9	22.6	23.1	23.4	23.7	23.8	23.9
ES	nRMSE	22.4	30.3	33.5	35.5	36.9	37.0	36.6	34.4	36.6	36.9
	nMAE	13.3	19.3	21.5	23.3	24.1	24.4	24.5	24.6	25.0	25.0
ARTU R = 0	nRMSE	22.2	30.8	30.7	32.4	33.5	34.0	34.3	34.4	34.7	34.8
	nMAE	14.8	21.3	20.9	27.8	22.7	23.2	23.5	23.6	23.8	23.9
ARTU R = 0.01	nRMSE	21.3	27.6	30.3	32.3	33.4	33.9	34.2	34.4	34.7	34.8
	nMAE	13.9	18.7	20.4	21.7	22.5	23.0	23.3	23.6	23.8	23.8
ARTU R = 0.05	nRMSE	21.3	27.5	30.3	32.3	33.4	33.9	34.2	34.5	34.7	34.8
	nMAE	13.9	18.6	20.4	21.7	22.5	23.0	23.3	23.6	23.8	23.8
ARTU R = 0.1	nRMSE	21.3	27.5	30.3	32.3	33.4	33.9	34.2	34.5	34.7	34.8
	nMAE	13.9	18.6	20.4	21.7	22.5	23.0	23.3	23.7	23.8	23.8
COMB	nRMSE	20.5	26.9	30.8	33.9	35.9	36.5	36.5	36.2	36.3	36.3
	nMAE	12.4	17.2	19.5	21.3	22.6	23.2	23.6	23.8	24.0	24.0

Table 2
MASE for the six different benchmark methods at Ajaccio, France. The lowest value is highlighted in bold.

	PER	CLIM	CLIPER	ES	ARTU R = 0	ARTU R = 0.01	ARTU R = 0.05	ARTU R = 0.1	COMB
MASE	52.10	124.11	44.45	46.40	45.23	44.29	44.27	44.28	43.64

we concentrate on the case where $R = 0.05$ since it is the one which proposes the lowest MASE. The combination of the models is undoubtedly the best alternative for short horizons whereas the ARTU model is the one for larger horizons. The results of these two models is given in Fig. 3 where a very good similarity with measurements is visible even during the most difficult winter months from a forecasting point of view. As conclusion, within the scope of this series of simulations, the proposal to use COMB as a baseline seems appropriate considering all the horizons.

4.2. Multi-site study

The results of the different forecasting methods are shown for four sites, one of which is relatively close to the previous one at about 105 km (Bastia). The other sites have entirely different climatological characteristics (Melbourne, Le Raizet and Nancy). The characteristics of these sites are shown in Table 3.

As can be seen from Table 4, the conclusions stated in the previous Section remain valid. Climatology-persistence systematically improves the persistence and is itself systematically improved by the ARTU method. The combination of these methods (COMB) remains the best alternative according the results.

If we focus on the site with the lowest forecastability (Nancy), Fig. 4 shows the evolution of the nRMSE and nMAE as function of the forecast horizon and confirms that in this particular case the ES model is the most suitable. This surprising result is probably related to the fact that for this site the clear sky model is less efficient than for the others. Concerning the site with the lowest forecast errors (Melbourne), we see in Fig. 4 that COMB has good performance compared with the other models mainly over shorter horizons. Drawing conclusions from this series of simulations, COMB and ARTU are the methodologies that give the best reliability.

4.3. Sensitivity to change in data type

In this Section, we study the sampling frequency of solar

irradiance measurements. Moreover, here, we focus on inclined (or tilted) radiation (TGI of 30°) [41], which is generally more difficult to predict because the anisotropy of the sky diffusion plays important role and is difficult to quantify. No hypothesis for the nature of radiation has been formulated in Section 2; it is therefore a question of checking if the preceding conclusions are valid in this particular case. The data used were measured on the Tilos site, a small Greek island with a Mediterranean climate (36°26'00"N, 27°22'00"E, alt. 100 m). This site has a forecastability $F = 82.5$; the data were acquired every 15 min and the climate of the site has a KG of Csa type. For the MASE calculation, m (the period in Eq. (17)) was modified and taken equal to 52 ($= 13 \times 4$) because there are 4 measurements per hour. The results are available in Table 5.

In this specific case, the fact that the measurements are so close together dramatically changes the results. Persistence is by far the best model for the 1-h horizon but it becomes the worst as the horizon increases. The best compromise seems to be COMB, although relatively bad at 150 min. In operations, it is not uncommon to study time steps of 15 min, particularly for piloting solar energy station (power or energy management system) but not up to a horizon of 150 min. In this case it will be necessary to use hourly data. Based on this observation, the COMB model is the one that provides more reliability.

4.4. Concerning other meteorological data

This section is dedicated to the study of meteorological time series other than solar. These are hourly series of ambient temperature [42] and 10-m wind speeds [43]. Temperature is largely periodic and similar what we observed with solar radiation, though wind speed is entirely different and exhibits little difference between day and night. The chosen site is Nancy, already studied in Section 4.2. The MASE evaluation is calculated in this case with its non-periodic version. Here we quantify the importance of the stationarity of the data. Table 6 presents the temperature results related to the different models assuming $I_{CS} = 1(\forall t)$. This

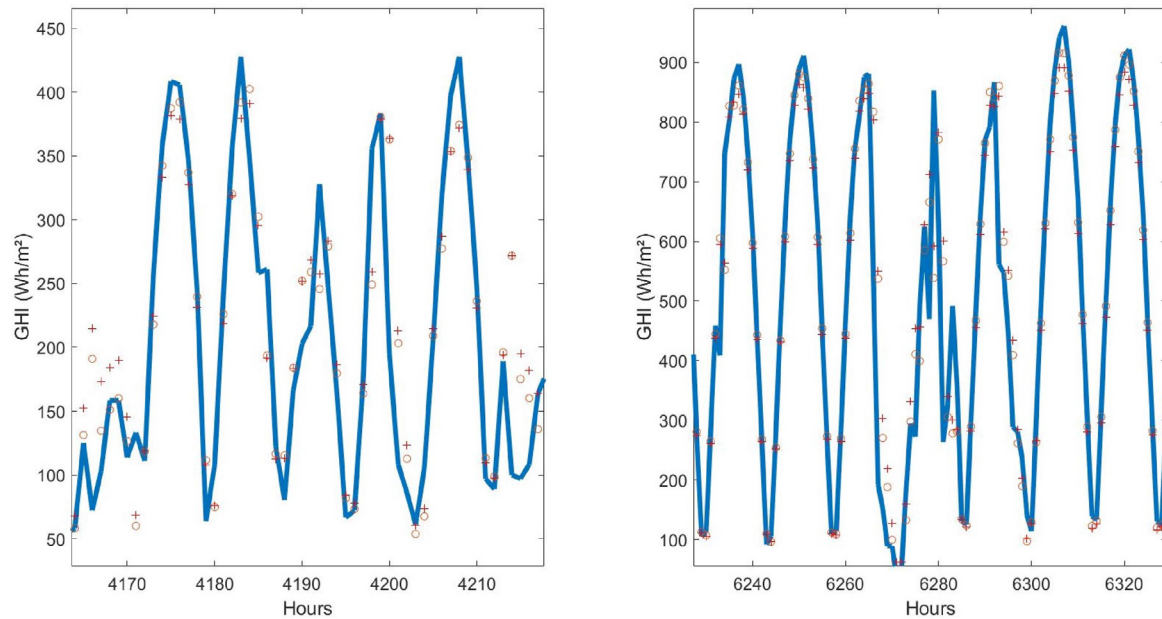


Fig. 3. Prediction (1 h horizon) for Ajaccio in winter (left) and summer (right). The blue lines correspond to the measurements (irradiation), the yellow circles to COMB, and red crosses to ARTU with $R = 0.05$.

Table 3

Characteristics of the studied sites. Climate classification KG according to Ascencio-Vasquez et al. [34] and Forecastability F according to Voyant et al. [35].

Name	Localisation	Coordinates	Alt (m)	KG	$F(\%)$
Ajaccio	France (Corsica)	41°55' 36" N 08°44' 13" E	30	Csa	68
Bastia	France (Corsica)	42°39' 14" N 09°39' 59" E	30	Csa	62.1
Nancy	France (Metropolitan)	48°41' 31" N 06°11' 03" E	271	Cfb	50.2
Le Raizet	France (Guadeloupe)	16°16' 15" N 61°30' 16" W	11	Af	58.2
Tilos	Greece	36°26' 00" N 27°22' 00" E	100	Csa	82.5
Melbourne	Australia	37°48' 50" S 144°57' 47" E	31	Cfb	63.2

Table 4

Comparison of the 4 sites and results (MASE) for Ajaccio (see Section 4.1). Error computed for all horizons comprise between 1 h and 10 h.

	PER	CLIPER	ES	ARTU $R = 0.05$	COMB
Ajaccio	52.10	44.45	46.40	44.27	43.64
Bastia	67.28	52.95	54.66	52.72	52.36
Nancy	59.85	59.25	54.06	58.49	54.23
La Raizet	70.60	51.76	55.03	51.52	52.77
Melbourne	66.12	56.70	57.20	56.37	54.52

corresponds to the non-stationary case in which x_t and κ_t are considered equal. We can expect that all models are penalized from the absence of deseasonalization of the time series, which particularly impacts ARTU.

The results do not correspond to what we have observed so far. The dynamic range of the signal tested is very low and the regularity of the successive measurements make it a good candidate for the use of persistence. Although the results are not good for the very short-term horizons; ARTU remains the method which is the most efficient on average. However, given its simplicity and the very good results observed in the first horizons, ES is undoubtedly

the reference method that should be used to properly characterize temperature forecasts. This suggests that as soon as the signal is regular with low variability, the forecast reference must imperatively be made with exponential smoothing. If we now focus on the estimation of wind speeds (Table 7 still considering $I_{CS} = 1$), the conclusions are different yet not inconsistent given the large difference in variability between these two meteorological quantities.

We observed that the climatology-persistence performed well in the case of solar radiation and its prediction. We see here that this is the best of the tested models and undoubtedly the easiest to set up. The ARTU model gives fairly comparable results. If, in the case of temperature, the ACF values were close from one lag to another, here it is the reverse; the correlations become insignificantly different from 0 very quickly. We observe an exponential decay (see Section 2.2.1), which suggests that the best model is indeed an AR(1) and therefore it is not surprising that CLIPER performs best and that ARTU does not bring a real added value to this case.

Using I_{CS} as defined in Section Appendix C with the time decomposition, we are able to see the impact of seasonal adjustment on the results. Tables 8 and 9 show the results according the temperature and the wind speed, respectively.

Even though this result has been understood for some time, we measure the importance of the ratio to trend and of the seasonal adjustment with these 2 Tables. The temperature results are very good for all methods and particularly for ARTU ($nMAE = 9.96$ for a 10-h horizon). It is likely that few machine-learning methods can significantly improve this result. For wind speeds, CLIPER remains the best way to make simple forecasts, even if ARTU and COMB can be very interesting alternatives, but also more complicated to set up. As in the previous sub-sections, a conclusion to propose COMB and ARTU as the best compromise of reference methods can be made based on the results of this section.

The field of wind forecasting is a very particular and complicated discipline. The choice to use the mean to try to make the series stationary is simple to establish though is perhaps not the best approach should we seek better results. Though not mentioned here, single, double or seasonal differencing would all have the effect of improving stationarity.

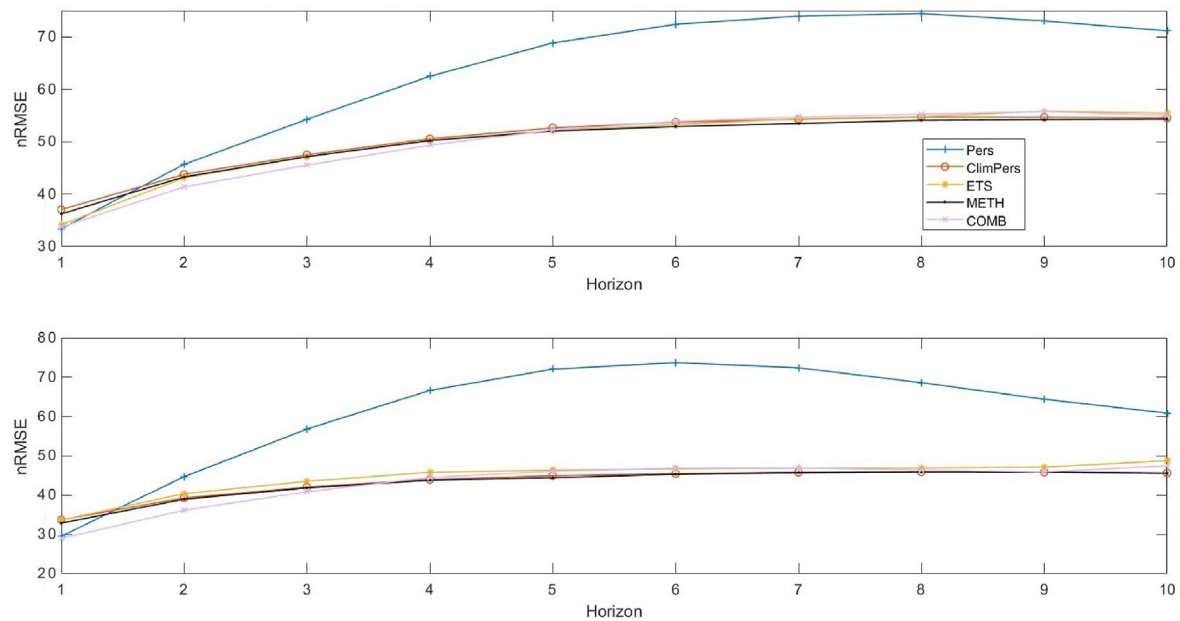


Fig. 4. nRMSE (in %) evolution as function of the forecast horizon for Nancy (top) and Melbourne (bottom).

Table 5

Error metrics from the Tilos evaluation. Global tilted irradiance measured every 15 min.

Horizon	Metrics	PER	CLIPER	ES	ARTU $R = 0.05$	COMB
15 min	nRMSE	13.52	19.27	24.86	18.71	14.00
	nMAE	6.13	12.16	13.97	11.75	8.26
75 min	nRMSE	58.87	24.99	38.36	24.89	26.98
	nMAE	20.91	15.83	23.15	15.73	15.34
150 min	nRMSE	116.9	25.98	37.55	25.95	40.17
	nMAE	36.00	16.39	23.71	16.38	19.91
	MASE	48.49	33.04	47.52	32.61	33.04

Table 6

Error metrics from the temperature evaluation without seasonal adjustments.

Horizon	Metrics	PER	CLIPER	ES	ARTU $R = 0.05$	COMB
15 min	nRMSE	8.90	8.85	9.03	17.16	8.79
	nMAE	6.32	6.26	6.45	13.59	6.35
75 min	nRMSE	27.91	26.21	27.16	26.25	26.13
	nMAE	21.90	20.36	21.60	20.07	20.30
150 min	nRMSE	39.28	34.36	29.36	30.67	31.45
	nMAE	33.62	28.17	25.09	24.80	26.39
	MASE	352.4	310.9	312.5	290.9	307.2

Table 7

Error metrics concerning the wind speed evaluation without seasonal adjustments.

Horizon	Metrics	PER	CLIPER	ES	ARTU $R = 0.05$	COMB
15 min	nRMSE	46.47	41.78	44.28	47.00	42.33
	nMAE	31.98	30.18	32.19	34.62	30.47
75 min	nRMSE	70.23	52.67	53.59	52.65	54.22
	nMAE	53.22	39.42	40.12	39.39	40.85
150 min	nRMSE	73.73	53.04	53.96	53.16	55.10
	nMAE	56.31	39.53	40.49	39.58	41.50
	MASE	159.6	120.2	122.9	120.2	124.5

5. Conclusions

In this paper, we proposed a new way of evaluating forecasts made in the field of meteorology and more specifically on solar radiation (Statistical Reference Method i.e. SRM). As the current

Table 8

Error metrics concerning the temperature evaluation with seasonal adjustments.

Horizon	Metrics	PER	CLIPER	ES	ARTU $R = 0.05$	COMB
15 min	nRMSE	6.08	5.91	6.19	6.85	6.02
	nMAE	4.01	4.03	4.17	5.02	4.06
75 min	nRMSE	14.66	12.19	13.86	12.17	12.82
	nMAE	10.25	9.06	9.82	9.06	9.23
150 min	nRMSE	19.62	14.07	14.66	13.12	14.44
	nMAE	14.01	10.74	11.01	9.96	10.87
	MASE	149.5	131.8	134.5	129.3	130.8

Table 9

Error metrics concerning the wind speed evaluation with seasonal adjustments.

Horizon	Metrics	PER	CLIPER	ES	ARTU $R = 0.05$	COMB
15 min	nRMSE	39.87	35.11	36.30	34.95	34.90
	nMAE	28.64	25.41	26.14	25.30	25.18
75 min	nRMSE	55.21	41.32	42.24	41.33	42.10
	nMAE	39.91	29.70	30.59	29.69	30.51
150 min	nRMSE	58.53	42.07	42.85	42.26	43.31
	nMAE	42.79	30.09	30.84	30.21	31.30
	MASE	118.8	90.0	91.4	89.7	91.7

practice is to compare the results of elaborate forecasting methods with those of simple methods (naive or reference), it should be appreciated that simple methods can evolve and improve. In the solar world for example, the default reference model is to use persistence and filtration, which in the end is quite similar to an AR(2) with the additional quality that no learning process is necessary. We title this new method ARTU which is the main innovation of this work.

In this paper, we evaluated all reference methods (PER, CLIM, CLIPER and ES) under the same formalism based on putting forward the measurement error, seasonality and forecast error through covariance estimation. This resulted in an evolution of the technique and the proposition of a new, simple model that mixes prediction and filtration, which in the end is quite similar to an AR(2) with the additional quality that no learning process is necessary. We title this new method ARTU which is the main innovation of this work.

The use of the classical exponential smoothing (ES) tools has been demonstrated to be of use again. Though it has been used in

some renewable energy papers [44], it is not included as part of the references nor of the naive methods. However, we have shown in this paper that ES belongs in that category.

The choice to consider combination (COMB) as a benchmark method can be seen as a “curiosity”, but in conclusion, when we combine two (or more) benchmarks, the result meets the criteria to be considered a benchmark method. Averaging makes it possible to overcome outliers and smooth forecasts. After this study, we can recommend the use of the combination of simple models such as PER, CLIPER, ES and ARTU. The main rules and recommendations of this study are:

- ↗ Reference models are critical to justify any new forecasting approach;
- ↗ The most appropriate benchmarking method depends on the nature of the variable being forecasted (seasonality, periodicity, forecastability) as well as the forecast horizon;
- ↗ The use of exponential smoothing for the estimation of “regular” series (high forecastability) should not be overlooked;
- ↗ The use of one error metrics specific to the time series forecasting community (MASE) are real assets deserving of uptake in more applied fields like solar energy engineering;
- ↗ The forecaster should endeavor to discover the benchmarking method most appropriate for their needs by either using them all (PER, CLIPER, ES, ARTU and COMB) and taking the best approach or by justifying the selection;
- ↗ In conclusion, COMB appears to be the best method among all those tested.

The development of increasingly sophisticated forecasting methods is necessary, but it is crucial to be able to evaluate them and to benchmark their efforts against the expected results. It is important to offer the most efficient benchmark in order to be able to compare studies with each other and identify the most efficient models even if it comes down to a fine detail. It is important to mention that in the solar energy community, it is usual to consider that an improvement of 1% in the prediction is approximately equivalent to an increase of 2% in economical gain [12]. For very large installations the gain can quickly become substantial.

CRedit authorship contribution statement

Cyril Voyant: Conceptualization, Project administration, Formal analysis, Writing – original draft, Supervision. **Gilles Notton:** Conceptualization, Writing – review & editing, Supervision. **Jean-Laurent Duchaud:** Writing – review & editing, Visualization. **Luis Antonio García Gutiérrez:** Writing – review & editing, Visualization. **Jamie M. Bright:** Writing – review & editing, Supervision, Investigation. **Dazhi Yang:** Writing – review & editing, Supervision, Formal analysis, Investigation, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Proof of CLIPER

From Eq. (4):

$$\mathbb{E}[x_{t+h}^c x_t^c] = \mathbb{E}[\alpha x_t^c x_t^c + \omega_{t+h} x_t^c], \quad (\text{A.1})$$

and from the definition of a white noise (ω uncorrelated with other variables) and of a stationary process (x^c), it becomes $\mathbb{E}[\omega_{t+h} x_t^c] = 0$.

Then let σ_x^2 and $\gamma(h)$ be, respectively, the variance of x_t (equal to the variance of x_t^c) and the autocovariance factors [45] between x_t^c and x_{t+h}^c ,² we observe that $\mathbb{E}[x_t^c x_t^c] = \sigma_x^2$ and as x_t^c is a centered variable, that $\mathbb{E}[x_t^c x_{t+h}^c] = \gamma(x_t^c, x_{t+h}^c) + (\mathbb{E}[x_t^c])^2 \equiv \gamma(h)$. Hence, Eq. (4) can be modified and simplified as ($\mathbb{E}[z_{t+h}] = \mathbb{E}[z_t]$):

$$\gamma(h) = \mathbb{E}[\alpha x_t^c x_t^c] + \mathbb{E}[\omega_{t+h} x_t^c] = \alpha \sigma_x^2. \quad (\text{A.2})$$

This implies that there is a simple link between α and the autocorrelation factor ρ (defined from the ratio between the autocovariance and the variance):

$$\alpha = \gamma(h) / \sigma_x^2 = \rho(h). \quad (\text{A.3})$$

Notice that considering the AR(1) model in Eq. (1) and the Yule–Walker equations or using Eq.(A.3) one can conclude that the prediction is now related to:

$$\hat{x}_{t+h}^c = \rho(h) x_t^c, \quad (\text{A.4})$$

as $x_t^c = x_t - \mathbb{E}[x]$ and $\mathbb{E}[x] = \mathbb{E}[y]$, Eq.(A.4) can be replaced with:

$$\hat{x}_{t+h} = \rho(h)[x_t - \mathbb{E}(x)] + \mathbb{E}(x) = \rho(h)[x_t - \mathbb{E}(y)] + \mathbb{E}(y). \quad (\text{A.5})$$

Let's not forget that x_t is an unknown quantity but that it is possible to have a measure of it through the variable called y_t (see Eq. (2)), this leads us to the following result:

$$\hat{x}_{t+h} = \rho(h) y_t + [1 - \rho(h)] \mathbb{E}(y). \quad (\text{A.6})$$

From a classical point of view, this equation is no different than the forecast by an AR(1) model using Yule–Walker equations, in the case of solar radiation prediction, this predictor (convex combination³ between climatology and persistence) is a particular reference called “climatology–persistence combination”. It tends to become the reference model in comparative studies [19].

Note that by taking the expectation of Eq. (1) multiplied by x_{t+h}^c and because $\mathbb{E}[\omega_t x_t^c] = \sigma_\omega^2$, it becomes:

$$\sigma_x^2 = \alpha \gamma(h) + \sigma_\omega^2. \quad (\text{A.7})$$

Replacing α with $\rho(h)$ in Eq.(A.7), we obtain:

$$\sigma_\omega^2 = \sigma_x^2 [1 - \rho(h)^2], \quad (\text{A.8})$$

where σ_ω^2 corresponds to the forecast error estimated by the Euclidean norm (mean square error; MSE) in the unbiased case. Because the correlation coefficient $\rho(h)$ is between 0 and 1, σ_ω^2 can take values between σ_x^2 ($\hat{x}_{t+h} = \bar{y}$, see Section 2.1.2) and 0 ($\hat{x}_{t+h} = y_t$, see Section 2.1.1). For example (see Appendix D.3 for details), when $\rho(h)$ takes the value of 0.7, the model is written $\hat{x}_{t+h} = 0.7 y_t + 0.3 \mathbb{E}(y)$ and the expected error is about $0.5 \sigma_x^2$. Of course, this theoretical error is not observed in reality, since we make the assumption that the observed phenomenon can be modeled by an AR(1), which is not necessarily true and which thus induces an additional part of uncertainty and an increase in the observed error. Moreover, the measurement error, represented by the parameter σ_v^2 in Eq. (2), also penalizes the forecast error.

² or x_{t-h}^c because autocorrelation is an even function while $x_t^c \in \mathbb{R}$ is a wide-sense stationary process.

³ In convex geometry, a convex combination is a linear combination of points in an affine space where all coefficients are non-negative and sum to 1.

Appendix B. Proof of ARTU

Equations (10)–(15) allow one to understand the link that can exist between a classical AR(2) and the method that we propose in this section: the prediction (\hat{x}_{t+h}) depends on two previous measures (y_t and y_{t-h}). The goal of this method is to find a mathematical formulation for α and K . A classic way to do this is to minimize the mean square error (MSE) which takes the form:

$$\text{MSE} = \mathbb{E}[(x_{t+h}^c - \hat{x}_{t+h}^c)^2] = \mathbb{E}[(x_{t+h}^c - \alpha x_t^c - K(y_t^c - \hat{x}_t^c))^2]. \quad (\text{B.1})$$

From Eq. (2), considering v_t (variance σ_v^2) has a covariance with all random variables null, Eq.(B.1) can be replaced with Eq.(B.2). This simplification is possible under the assumption of orthogonality (or uncorrelation working with centered series).

$$J = \begin{pmatrix} \alpha^2 - 2\rho(h)\alpha + 1 + R & \rho(2h) + 2K\alpha - 2K\rho(h) - 2\alpha\rho(h) + 1 \\ \rho(2h) + 2K\alpha - 2K\rho(h) - 2\alpha\rho(h) + 1 & K^2 - 2K\rho(h) + 1 \end{pmatrix}. \quad (\text{B.7})$$

$$\text{MSE} = K^2\sigma_v^2 + \mathbb{E}(u)^2 \text{ with } u = x_{t+h}^c - \alpha x_t^c - x^c z_t + K\alpha x_{t-h}^c. \quad (\text{B.2})$$

A sufficient condition to find the optimal values of α and K consists of finding their values that minimize the MSE, by equating to 0 its gradient. This is equivalent to solving $\partial\text{MSE}/\partial K = 0$ (see Eq. B.3) and $\partial\text{MSE}/\partial\alpha = 0$ (see Eq. B.4). Note that interchanging the derivative with expectation can be done using the dominated convergence theorem.⁴

$$\begin{aligned} \frac{\partial\text{MSE}}{\partial K} &= 2K\sigma_v^2 + \frac{\partial\mathbb{E}(u)^2}{\partial K} = 2K\sigma_v^2 + \mathbb{E}\left(\frac{\partial u^2}{\partial u} \frac{\partial u}{\partial K}\right) \\ &= 2K\sigma_v^2 + 2\mathbb{E}[u(\alpha x_{t-h}^c - x_t^c)] = 0. \end{aligned} \quad (\text{B.3})$$

An identical reasoning allows us to establish Eq.(B.3) and Eq.(B.4) concerning the derivative with respect to K and α .

$$\frac{\partial\text{MSE}}{\partial\alpha} = 2\mathbb{E}[u(Kx_{t-h}^c - x_t^c)] = 0. \quad (\text{B.4})$$

After some mathematical simplifications, the solution of the problem amounts to finding α and K solutions of the system described by Eq.(B.5). It is a system of quadratic equations with 2 unknowns (K and α) of degree 2.

$$\begin{cases} K(\sigma_v^2 + \sigma_x^2) + \alpha(\gamma(2h) + \sigma_x^2) - 2K\alpha\gamma(h) - \alpha^2\gamma(h) + K\alpha^2\sigma_x^2 = \gamma(h) \\ K(\gamma(2h) + \sigma_x^2) - 2K\alpha\gamma(h) + \alpha\sigma_x^2 - K^2\gamma(h) + K^2\alpha\sigma_x^2 = \gamma(h) \end{cases} \quad (\text{B.5})$$

Solving this system is not trivial, and it is best to slightly modify it to make the task easier. The more convenient form is certainly obtained dividing the two equations by σ_x^2 ($\in \mathbb{R}^*$) and defining a

new variable $R = \sigma_v^2/\sigma_x^2$. A discussion about it is proposed in Section 2.2.1. This modification allows to highlight the auto-correlation coefficients ρ as shown in Eq.(B.6). $\rho(h)$ is related to the correlation between x_t^c and itself delayed by h lags, while $\rho(2h)$ concerns a delay of $2h$ lags.

$$\begin{cases} K(1+R) + \alpha[1 + \rho(2h)] - 2K\alpha\rho(h) - \alpha^2\rho(h) + K\alpha^2 = \rho(h) \\ K[1 + \rho(2h)] - 2K\alpha\rho(h) + \alpha - K^2\rho(h) + K^2\alpha = \rho(h) \end{cases} \quad (\text{B.6})$$

Among the five couples of solutions, we retain only those that have a physical reality: the real ones (most of the time there are 2). Rather than finding exact and symbolic solutions since there are no simple ones (according to our knowledge and what we could find in the literature), we propose to make them explicit by means of Levenberg–Marquardt algorithm [46]. We used 100 iterations with random initializations of the 2 unknowns (K and α) between -1 and 1 . This method requires at each iteration to compute $(J^T J + \lambda \text{diag}(J^T J))^{-1} J^T$ where the Jacobian (J) is defined in Eq.(B.6) such as:

The Monge theorem allows to retain only the critical points compatible with a local minimum of the MSE. In mathematical analysis, this theorem is used to study the behavior of a function with two variables (K, α) in the neighborhood of a critical point (K^*, α^*). The retained solutions satisfy the condition of the local extremity mentioned in Eq.(B.8) and of strict local minimum point exposed in Eq.(B.9).

$$\left[\frac{\partial^2\text{MSE}}{\partial K \partial \alpha}(K^*, \alpha^*) \right]^2 - \frac{\partial^2\text{MSE}}{\partial K^2}(K^*, \alpha^*) \frac{\partial^2\text{MSE}}{\partial \alpha^2}(K^*, \alpha^*) < 0, \quad (\text{B.8})$$

$$\frac{\partial^2\text{MSE}}{\partial K^2}(K^*, \alpha^*) > 0. \quad (\text{B.9})$$

In case there is more than one solution (according to a particular triplet $[R, \rho(h), \rho(2h)]$), we choose to calculate the associated MSE and to retain only those which minimize this quantity. This can be done from the two partial derivatives defined in Eq.(B.6) and respectively in Eq.(B.10) and Eq.(B.11) concerning K and α .

$$\frac{\partial\text{MSE}}{\partial K} = K(1+R) + \alpha[1 + \rho(2h)] - 2K\alpha\rho(h) - \alpha^2\rho(h) + K\alpha^2 - \rho(h), \quad (\text{B.10})$$

$$\frac{\partial\text{MSE}}{\partial \alpha} = K[1 + \rho(2h)] - 2K\alpha\rho(h) + \alpha - K^2\rho(h) + K^2\alpha - \rho(h). \quad (\text{B.11})$$

A succession of integration (with respect to K then with respect to α) allows to determinate without too much difficulty a function (MSE in our case) defined by its partial derivatives. The result is given in Eq.(B.12) with a constant of integration c ($\in \mathbb{R}$).

⁴ It is one of the main theorems of Lebesgue's integration theory giving a sufficient condition for the convergence of expected values of random variables.

$$\begin{aligned} \text{MSE} = K^2 \left[\frac{R}{2} + \frac{1}{2} \right] - K\rho(h) - \alpha[K^2\rho(h) - K[\rho(2h) + 1] + \rho(h)] \\ + \alpha^2 \left[\frac{K^2}{2} - K\rho(h) + \frac{1}{2} \right] + c \end{aligned} \quad (\text{B.12})$$

Appendix C. Seasonality and Stationarity

The methods presented in Section 2 are based on the assumption that the time series studied are non-seasonal and devoid of trend. The latter hypothesis is always the case in meteorology on short time scales and particularly in solar radiation (stability of the Holocene climate) that has no significant inter-annual change (the annual average of the signal can be considered constant over a period of 10 years). The first hypothesis on the other hand is by nature invalid for weather series. Fortunately in solar energy forecasting, a transformation for removing the seasonality can be calculated very easily, via a clear-sky model [26]. Clear-sky irradiance (I_{CS}) is the solar radiation incident on a horizontal surface under a cloud-free sky. Thus, the global horizontal irradiance (I_{GH}) are related to the seasonally adjusted variable, namely, the clear-sky index (κ), through the following:

$$\kappa \equiv \kappa = I_{GH}/I_{CS}, \quad (\text{C.1})$$

If I_{CS} is well-modeled, in theory x_t (the clear sky index) and y_t (its measurement) are without seasonality and are bounded between 0 and somewhere between 1 and 1.5—the upper bound would depend on the cloud meteorology; in practice, an upper bound of 1.2 is often used [47]. However, in terms of seasonality, it has been shown that even the best clear-sky models today are unable to completely remove it, resulting in a nonstationary clear-sky index time series [48]. In practice, I_{CS} is often calculated from the Lambert–Beer type relations [49] or using directly the data from CAMS McClear service [50].

In the case of studying temperature and wind speed, the methodology is equivalent. I_{CS} is replaced by the measure of the meteorological quantity by calculating for each hour of the year the average of the same hours concerning the previous years (model free). Generally, this procedure behaves as a particular kind of low-pass filter, and I_{CS} is equivalent to smoothed series of temperature and wind speed.

The impact of seasonal adjustment can be quantified from Eq.(C.2) and the statistic $t(m)$ [51] or computing the squared m -th auto-correlation of the series and comparing it to a $\chi(1)^2$ distribution as described in Ref. [52].

$$t(m) = q_{1-\alpha/2} \sqrt{\frac{1 + 2 \sum_{i=1}^{m-1} \rho^2(i)}{n}}, \quad (\text{C.2})$$

where q is the quantile function of the standard normal distribution and $100(1 - \alpha)\%$ corresponds to the confidence level; a 90% confidence level is often used. m is the number of the periods within a seasonal cycle (for example, 24 and $8760 = 24 \times 365$ for hourly data). So, the larger the value of $t(m)$, the larger the seasonality is. If $|\rho(m)| < t(m)$, the series can be considered deseasonalized. However, one must be careful because, like all statistical tests, this test is very sensitive to the size of the sample, so it is more relevant to subsample (randomly) the data if one wants a better interpretation of the test. We can assume $n = 100$ without loss of generality.

It is important to note that the forecasting of solar radiation time series is a special topic. The models presented previously and simulated in the next Section do not refer to a consideration of seasonality because of the clear-sky model. Forecasting the clear-sky index, and using the Error–Trend–Seasonal (ETS, N = none, A = additive) framework terminology of Hyndman and Khandakar [53], PER and CLIM are (N,N,N), CLIPER, ES, ARTU and COMB are (A,N,N). None of them require an optimization and training phases and some of them can be used with only recent measurements.

Appendix D. Algorithms

The pseudo-codes detailed in this section are adapted to the case of global irradiance ($I_{GH}(t)$) with $t \in \{1, \dots, T^*, \dots, T^{\text{OutSample}}\}$ though they can be modified for any kind of time series. Up to now, we deliberately neglect the phenomena of over irradiance ($I_{GH}(t) \in [0, I_{CS}(t)]$), however depending on the time step, the clear sky model used and the quality of the time-stamp, it could be necessary to multiply I_{CS} by an arbitrary coefficient β (generally between 1 and 2). Anyway, all models must make it possible to provide forecasts for all hours of the day and night, however the validation of the results is only performed on the daytime hours (authorizing solar elevation greater than $5 - 10^\circ$). Even if it is not the purpose of this study, it is important not to neglect the forecasts of the first and last hours of the daylight, they can be very important for energy management systems. Often, the real reasons for which a filtration is operated because of the poor quality of the detection concerning these hours and the strong repercussions (periodic peaks on κ) that a poor time-stamp can induce.

Appendix D.1. Persistence

This persistence predictor (Algorithm 1) is certainly the simplest method use in order to operate predictions with reliability.

Algorithm 1. PER.

Algorithm 1 PER

Input: $I_{CS}, I_{GH}, h > 0, \beta \in [1, 2]$

Output: $\hat{I}_{GH}(t+h)$ with $t \in [T^*, T]$

$n \leftarrow 0$

repeat

$n \leftarrow n + 1$

until $I_{CS}(t-n) \neq 0$

$\text{Pred} \leftarrow \min(I_{GH}(t-n) \times I_{CS}(t+h)/I_{CS}(t-n), \beta \times I_{CS}(t+h))$

return Pred

Appendix D.2. Climatology

Even if this predictor ([Algorithm 2](#)) is never used in practice, it is an important way to gauge results in solar prediction study. When no model of knowledge is available, a moving average can be a good alternative. One of the characteristics of this model is that the observed forecast error is constant whatever the horizon considered. The filtering parameter (ϵ) is usually taken close to 10 (Wh/m²) while certain authors prefer use a threshold between 5° and 10° concerning the solar elevation. The information linked to the cloudiness being observable only in the presence of daylight, only these moments must be used. This means that at sun-up, it is the data from the day before that is used, so we understand the limit of statistical forecast models using only endogenous quantities.

Algorithm 2. CLIM.**Algorithm 2** CLIM**Input:** $I_{CS}, I_{GH}, h > 0, \epsilon \in [1, 30]$ **Output:** $\hat{I}_{GH}(t+h)$ with $t \in [T^*, T]$ for $n := 1$ to T^* doif $I_{CS}(n) < \epsilon$ then $\kappa(n) = \emptyset$

else

 $\kappa(n) \leftarrow I_{GH}(n)/I_{CS}(n)$

end if

end for

 $\bar{\kappa} \leftarrow \text{mean}(\kappa(n))$ $\text{Pred}\kappa \leftarrow \bar{\kappa}$ $\text{Pred} \leftarrow \text{Pred}\kappa \times I_{CS}(t+h)$

return Pred

Appendix D.4. Exponential Smoothing

In practice, it is not necessary to calculate the smoothing on all the in-sample data, limiting to a range covering the daily periodicity ($max = 24$ h) or 2 times this ($max = 48$ h) is sufficient to obtain good results ([Algorithm 4](#)).

Appendix D.3. Climatology Persistence

CLIPER is undoubtedly the new standard of reference forecast for solar irradiation. As in the previous case ([Appendix D.2](#)), a filtering process is operated in [Algorithm 3](#) and the ϵ parameter is considered for this task.

Algorithm 3. CLIPER.**Algorithm 3** CLIPER**Input:** $I_{CS}, I_{GH}, h > 0, \beta \in [1, 2], \epsilon \in [1, 30]$ **Output:** $\hat{I}_{GH}(t+h)$ with $t \in [T^*, T]$ for $n := 1$ to T^* doif $I_{CS}(n) < \epsilon$ then $\kappa(n) = \emptyset$

else

 $\kappa(n) \leftarrow I_{GH}(n)/I_{CS}(n)$

end if

end for

 $\bar{\kappa} \leftarrow \text{mean}(\kappa(n))$ $\rho \leftarrow \text{ACF}(\kappa(n), \kappa(n-h))$ $nn \leftarrow 0$

repeat

 $nn \leftarrow nn + 1$ until $I_{CS}(t-nn) \geq \epsilon$ $\text{Pred}\kappa \leftarrow \min(\rho \times \kappa(t-nn) + (1-\rho) \times \bar{\kappa}, \beta)$ $\text{Pred} \leftarrow \text{Pred}\kappa \times I_{CS}(t+h)$

return Pred

Algorithm 4. ES.**Algorithm 4 ES**

Input: $I_{CS}, I_{GH}, h > 0, \beta \in [1, 2], \epsilon \in [1, 30], max \in [10 - 48]$
Output: $\hat{I}_{GH}(t + h)$ with $t \in [T^*, T]$

```

for  $n := 1$  to  $T^*$  do
  if  $I_{CS}(n) < \epsilon$  then
     $\kappa(n) = 1$ 
  else
     $\kappa(n) \leftarrow I_{GH}(n)/I_{CS}(n)$ 
  end if
end for
 $\bar{\kappa} \leftarrow \text{mean}(\kappa(n))$ 
 $\rho \leftarrow \text{ACF}(\kappa(n), \kappa(n - h))$ 
for  $nn := 0$  to  $max - 1$  do
  if  $I_{CS}(t - nn) < \epsilon$  then
     $\kappa(t - nn) = 1$ 
  else
     $\kappa(t - nn) \leftarrow I_{GH}(t - nn)/I_{CS}(t - nn)$ 
  end if
end for
Pred $\kappa \leftarrow \min(\rho \times \sum_{i=0}^{max-1} (1 - \rho)^i \times \kappa(t - i) + \bar{\kappa} \times (1 - \rho)^{max}, \beta)$ 
Pred  $\leftarrow \text{Pred}\kappa \times I_{CS}(t + h)$ 
return Pred

```

Appendix D.5. Proposed Methodology (ARTU)

In this version of the code (Algorithm 5), we propose to associate the night hours with a κ equal to 1 but another way which is slightly more complex but which gives very good results consists in neglecting the night hours by removing them completely as done in the Algorithm 3 (if $I_{CS}(n) < \epsilon$ then $\kappa(n) = \emptyset$). The method requires

knowledge of α and K , which is achieved by interpolating the $\mathcal{M}(R)$ matrices (see <https://github.com/cyrilvoyant/ARTU.git>). Knowing the correlation coefficients ($\rho(h)$ and $\rho(2h)$) and the measurement reliability ($R = 0, 0.01, 0.05, 0.1$) the interpolation allows an estimate of α and K for these three characteristic values.

Algorithm 5. ARTU.**Algorithm 5 ARTU**

Input: $I_{CS}, I_{GH}, h > 0, \beta \in [1, 2], \epsilon \in [1, 30], R \in [0, 0.01, 0.05, 0.1], \mathcal{M}(R)$
Output: $\hat{I}_{GH}(t + h)$ with $t \in [T^*, T]$

```

for  $n := 1$  to  $T^*$  do
  if  $I_{CS}(n) < \epsilon$  then
     $\kappa(n) = 1$ 
  else
     $\kappa(n) \leftarrow I_{GH}(n)/I_{CS}(n)$ 
  end if
end for
 $\bar{\kappa} \leftarrow \text{mean}(\kappa(n))$ 
 $\rho1 \leftarrow \text{ACF}(\kappa(n), \kappa(n - h))$ 
 $\rho2 \leftarrow \text{ACF}(\kappa(n), \kappa(n - 2h))$ 
 $(\alpha, K) \leftarrow \text{interpolate}(\mathcal{M}(R), \rho1, \rho2, R)$ 
 $S \leftarrow \alpha + K$ 
 $P \leftarrow \alpha \times K$ 
for  $nn := 0$  to  $h$  do
  if  $I_{CS}(t - nn) < \epsilon$  then
     $\kappa(t - nn) = 1$ 
  else
     $\kappa(t - nn) \leftarrow I_{GH}(t - nn)/I_{CS}(t - nn)$ 
  end if
end for
Pred $\kappa \leftarrow \min(S \times \kappa(t) - P \times \kappa(t - h) + (1 + P - S) \times \bar{\kappa}, \beta)$ 
Pred  $\leftarrow \text{Pred}\kappa \times I_{CS}(t + h)$ 
return Pred

```

References

- [1] P. Agathangelou, D. Trihinas, I. Katakis, A multi-factor analysis of forecasting methods: a study on the m4 competition, *Data* 5 (2020) 41, <https://doi.org/10.3390/data5020041>.
- [2] G. Notton, M.-L. Nivet, C. Voyant, C. Paoli, C. Darras, F. Motte, A. Fouilloy, Intermittent and stochastic character of renewable energy sources: consequences, cost of intermittence and benefit of forecasting, *Renew. Sustain. Energy Rev.* 87 (2018) 96–105, <https://doi.org/10.1016/j.rser.2018.02.007>. URL: <https://www.sciencedirect.com/science/article/pii/S1364032118300327>.
- [3] S. Makridakis, C. Fry, F. Petropoulos, E. Spiliotis, *The Future of Forecasting Competitions: Design Attributes and Principles*, 2021, 04879 arXiv:2102.
- [4] D. Gomez, A. Rojas, An empirical overview of the no free lunch theorem and its effect on real-world machine learning classification, *Neural Comput.* 28 (2016) 216–228, <https://doi.org/10.1162/NECO.a.00793>.
- [5] T. Gneiting, Making and evaluating point forecasts, *J. Am. Stat. Assoc.* 106 (2011) 746–762. URL: <http://www.jstor.org/stable/41416407>.
- [6] D. Yang, S. Alessandrini, J. Antonanzas, F. Antonanzas-Torres, V. Badescu, H.G. Beyer, R. Blaga, J. Boland, J.M. Bright, C.F. Coimbra, M. David, A. Frimane, C.A. Gueymard, T. Hong, M.J. Kay, S. Killinger, J. Kleissl, P. Lauret, E. Lorenz, D. van der Meer, M. Paulescu, R. Perez, O. Perpiñán-Lamigueiro, I.M. Peters, G. Reikard, D. Renné, Y.-M. Saint-Drenan, Y. Shuai, R. Urraca, H. Verbois, F. Vignola, C. Voyant, J. Zhang, Verification of deterministic solar forecasts, *Sol. Energy* 210 (2020) 20–37, <https://doi.org/10.1016/j.solener.2020.04.019>, special Issue on Grid Integration.
- [7] J.M. Bright, The impact of globally diverse ghi training data: evaluation through application of a simple Markov chain downscaling methodology, *J. Renew. Sustain. Energy* 11 (2019) 23703, <https://doi.org/10.1063/1.5085236>.
- [8] X. Yang, D. Yang, J.M. Bright, G.M. Yagli, P. Wang, On predictability of solar irradiance, *J. Renew. Sustain. Energy* 13 (2021), 056501, <https://doi.org/10.1063/5.0056918>.
- [9] E. Lorenz, J. Remund, S. Müller, W. Traunmüller, G. Steinmaurer, D. Pozo-Vazquez, J. Ruiz-Arias, V. Lara Fanego, L. Santigosa, M. Gastón, C. Kurz, L. Martín Pomares, C. Geijo, Benchmarking of Different Approaches to Forecast Solar Irradiance, 24th European Photovoltaic Solar Energy Conference, 2009.
- [10] M. Fliess, C. Join, C. Voyant, Prediction bands for solar energy: new short-term time series forecasting techniques, *Sol. Energy* 166 (2018) 519–528, <https://doi.org/10.1016/j.solener.2018.03.049>. URL: <https://www.sciencedirect.com/science/article/pii/S0038092X18302810>.
- [11] T. Soubdhan, J. Ndong, H. Ould-Baba, M.-T. Do, A robust forecasting framework based on the kalman filtering approach with a twofold parameter tuning procedure: application to solar and photovoltaic prediction, *Sol. Energy* 131 (2016) 246–259, <https://doi.org/10.1016/j.solener.2016.02.036>. URL: <https://www.sciencedirect.com/science/article/pii/S0038092X16001444>.
- [12] M. David, J. Boland, L. Cirocco, P. Lauret, C. Voyant, Value of deterministic day-ahead forecasts of pv generation in pv + storage operation for the Australian electricity market, *Sol. Energy* 224 (2021) 672–684, <https://doi.org/10.1016/j.solener.2021.06.011>. URL: <https://www.sciencedirect.com/science/article/pii/S0038092X21004862>.
- [13] J.-L. Duchaud, C. Voyant, A. Fouilloy, G. Notton, M.-L. Nivet, Trade-off between precision and resolution of a solar power forecasting algorithm for micro-grid optimal control, *Energies* 13 (2020). URL: <https://www.mdpi.com/1996-1073/13/14/3565>.
- [14] D. Yang, Standard of reference in operational day-ahead deterministic solar forecasting, *J. Renew. Sustain. Energy* 11 (2019) 53702, <https://doi.org/10.1063/1.5114985>. URL: .
- [15] D. Yang, A universal benchmarking method for probabilistic solar irradiance forecasting, *Sol. Energy* 184 (2019) 410–416, <https://doi.org/10.1016/j.solener.2019.04.018>. URL: <https://www.sciencedirect.com/science/article/pii/S0038092X19303457>.
- [16] A.H. Murphy, Climatology, persistence, and their linear combination as standards of reference in skill scores, *Weather Forecast.* 7 (1992) 692–698, [https://doi.org/10.1175/1520-0434\(1992\)007<0692:CPATLC>2.0.CO;2](https://doi.org/10.1175/1520-0434(1992)007<0692:CPATLC>2.0.CO;2).
- [17] T. Hong, P. Pinson, S. Fan, Global energy forecasting competition 2012, *Int. J. Forecast.* 30 (2014) 357–363, <https://doi.org/10.1016/j.ijforecast.2013.07.001>. URL: <https://www.sciencedirect.com/science/article/pii/S0169207013000745>.
- [18] T. Hong, P. Pinson, Y. Wang, R. Weron, D. Yang, H. Zareipour, Energy forecasting: a review and outlook, *IEEE Open Access J. Power Energy* 7 (2020) 376–388, <https://doi.org/10.1109/OAJPE.2020.3029979>.
- [19] D. Yang, Making reference solar forecasts with climatology, persistence, and their optimal convex combination, *Sol. Energy* 193 (2019) 981–985, <https://doi.org/10.1016/j.solener.2019.10.006>. URL: <https://www.sciencedirect.com/science/article/pii/S0038092X19309880>.
- [20] S. Makridakis, S. Wheelwright, H. Rob, in: *Forecasting: Methods and Applications*, third ed., Wiley, United States of America, 1997.
- [21] G. Bontempi, S. Ben Taieb, Y.-A. Le Borgne, Machine learning strategies for time series forecasting, in: M.-A. Aufaure, E. Zimányi (Eds.), *Business Intelligence: Second European Summer School, eBISS 2012, Brussels, Belgium, July 15–21, 2012, Tutorial Lectures*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, pp. 62–77, https://doi.org/10.1007/978-3-642-36318-4_3. URL: .
- [22] S.B. Taieb, R.J. Hyndman, Recursive and direct multi-step forecasting: the best of both worlds, Monash University, Department of Econometrics and business statistics, Tech. Rep. (2012).
- [23] Hui-Min Zhang, P. Duhamel, On the methods for solving yule-walker equations, *IEEE Trans. Signal Process.* 40 (1992) 2987–3000, <https://doi.org/10.1109/78.175742>.
- [24] C. Chatfield, The future of the time-series forecasting, URL: [https://doi.org/10.1016/0169-2070\(88\)90108-2](https://doi.org/10.1016/0169-2070(88)90108-2). Special Issue on: The Future of Forecasting, <https://www.sciencedirect.com/science/article/pii/S0169207088901082>.
- [25] J.G. De Gooijer, R.J. Hyndman, 25 years of time series forecasting, twenty five years of forecasting, *Int. J. Forecast.* 22 (2006) 443–473, <https://doi.org/10.1016/j.ijforecast.2006.01.001>. URL: <https://www.sciencedirect.com/science/article/pii/S0169207006000021>.
- [26] X. Sun, J.M. Bright, C.A. Gueymard, X. Bai, B. Acord, P. Wang, Worldwide performance assessment of 95 direct and diffuse clear-sky irradiance models using principal component analysis, *Renew. Sustain. Energy Rev.* 135 (2021), 110087, <https://doi.org/10.1016/j.rser.2020.110087>.
- [27] X. Sun, J.M. Bright, C.A. Gueymard, B. Acord, P. Wang, N.A. Engerer, Worldwide performance assessment of 75 global clear-sky irradiance models using principal component analysis, *Renew. Sustain. Energy Rev.* 111 (2019) 550–570, <https://doi.org/10.1016/j.rser.2019.04.006>.
- [28] J.M. Bright, X. Bai, Y. Zhang, X. Sun, B. Acord, P. Wang, irrady: Python package for merra-2 download, extraction and usage for clear-sky irradiance modelling, *Sol. Energy* 199 (2020) 685–693, <https://doi.org/10.1016/j.solener.2020.02.061>.
- [29] L. Vuilleumier, C. Félix, F. Vignola, P. Blanc, J. Badosa, A. Kazantzidis, B. Calpini, Performance evaluation of radiation sensors for the solar energy sector, *Meteorol. Z.* 26 (2017) 485–505, <https://doi.org/10.1127/metz/2017/0836>.
- [30] S. Makridakis, E. Spiliotis, V. Assimakopoulos, The m4 competition: results, findings, conclusion and way forward, *Int. J. Forecast.* 34 (2018), <https://doi.org/10.1016/j.ijforecast.2018.06.001>.
- [31] E. Spiliotis, H. Doukas, V. Assimakopoulos, F. Petropoulos, Chapter 4 - forecasting week-ahead hourly electricity prices in Belgium with statistical and machine learning methods, in: A. Dagoumas (Ed.), *Mathematical Modelling of Contemporary Electricity Markets*, Academic Press, 2021, pp. 59–74, <https://doi.org/10.1016/B978-0-12-821838-9.00005-0>.
- [32] F. Petropoulos, I. Svetunkov, A simple combination of univariate models, *Int. J. Forecast.* (2018), <https://doi.org/10.1016/j.ijforecast.2019.01.006>.
- [33] A. Timmermann, Chapter 4 forecast combinations, in: G. Elliott, C. Granger, A. Timmermann (Eds.), *Handbook of Economic Forecasting*, 1, Elsevier, 2006, pp. 135–196, [https://doi.org/10.1016/S1574-0706\(05\)01004-9](https://doi.org/10.1016/S1574-0706(05)01004-9). URL: <https://www.sciencedirect.com/science/article/pii/S1574070605010049>.
- [34] J. Ascencio-Vasquez, K. Brecl, M. Topic, Methodology of koppen-geiger-photovoltaic climate classification and implications to worldwide mapping of pv system performance, *Sol. Energy* 191 (2019) 672–685, <https://doi.org/10.1016/j.solener.2019.08.072>.
- [35] C. Voyant, P. Lauret, G. Notton, J.-L. Duchaud, A. Fouilloy, M. David, Z.M. Yaseen, T. Soubdhan, A Monte Carlo based solar radiation forecastability estimation, *J. Renew. Sustain. Energy* 13 (2021) 26501, <https://doi.org/10.1063/5.0042710>.
- [36] B. Espinar, P. Blanc, L. Wald, C. Hoyer-Klick, M. Schroedter-Homscheidt, T. Wanderer, On quality control procedures for solar radiation and meteorological measures, from subhourly to monthly average time periods, in: *European Geosciences Union 2012, Vienna, Austria, 2012*, p. 9134. URL: <http://adsabs.harvard.edu/abs/2012EGUGA..14.9134E>.
- [37] O. el Alani, H. Ghennoui, A. Ghennoui, Y.-M. Saint-Drenan, P. Blanc, N. Hanrieder, F.-E. Dahr, A visual support of standard procedures for solar radiation quality control, *Int. J. Renew. Energy Dev.* 10 (2021) 401–414, <https://doi.org/10.14710/ijred.2021.34806>.
- [38] R.J. Hyndman, A.B. Koehler, Another look at measures of forecast accuracy, *Int. J. Forecast.* 22 (2006) 679–688, <https://doi.org/10.1016/j.ijforecast.2006.03.001>.
- [39] D. Koutsandreas, E. Spiliotis, F. Petropoulos, V. Assimakopoulos, On the selection of forecasting accuracy measures, *J. Oper. Res. Soc.* (2021) 1–18, <https://doi.org/10.1080/01605682.2021.1892464>.
- [40] C. Voyant, J.G. De Gooijer, G. Notton, Periodic autoregressive forecasting of global solar irradiance without knowledge-based model implementation, *Sol. Energy* 174 (2018) 121–129, <https://doi.org/10.1016/j.solener.2018.08.076>. URL: <https://www.sciencedirect.com/science/article/pii/S0038092X18308466>.
- [41] M. David, P. Lauret, J. Boland, Evaluating tilted plane models for solar radiation using comprehensive testing procedures, at a southern hemisphere location, *Renew. Energy* 51 (2013) 124–131, <https://doi.org/10.1016/j.renene.2012.08.074>.
- [42] J. Corchado, C. Fyfe, Unsupervised neural method for temperature forecasting, *Artif. Intell. Eng.* 13 (1999) 351–357, [https://doi.org/10.1016/S0954-1810\(99\)00007-2](https://doi.org/10.1016/S0954-1810(99)00007-2). URL: <https://www.sciencedirect.com/science/article/pii/S0954181099000072>.
- [43] G. Giebel, J. Cline, H. Frank, W. Shaw, P. Pinson, B.-M. Hodge, G. Kariniotakis, J. Madsen, C. Möhrlen, Wind power forecasting: IEA wind task 36 & future research issues, *J. Phys. Conf.* 753 (2016) 32042, <https://doi.org/10.1088/1742-6596/753/3/032042>. URL: .
- [44] Z. Dong, D. Yang, T. Reindl, W.M. Walsh, Short-term solar irradiance forecasting using exponential smoothing state space model, *Energy* 55 (2013) 1104–1113, <https://doi.org/10.1016/j.energy.2013.04.027>.
- [45] S.G. Makridakis, in: *Forecasting: Methods and Applications*, third ed., Wiley, 1998.
- [46] J.J. Moré, *The levenberg-marquardt algorithm: implementation and theory*,

- in: G.A. Watson (Ed.), *Numerical Analysis*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1978, pp. 105–116.
- [47] H.T.C. Pedro, D.P. Larson, C.F.M. Coimbra, A comprehensive dataset for the accelerated development and benchmarking of solar forecasting methods, *J. Renew. Sustain. Energy* 11 (2019) 36102, <https://doi.org/10.1063/1.5094494>.
- [48] D. Yang, Choice of clear-sky model in solar forecasting, *J. Renew. Sustain. Energy* 12 (2020) 26101, <https://doi.org/10.1063/5.0003495>.
- [49] P. Ineichen, A broadband simplified version of the solis clear sky model, *Sol. Energy* 82 (2008) 758–762, <https://doi.org/10.1016/j.solener.2008.02.009>.
- [50] M. Lefèvre, A. Oumbe, P. Blanc, B. Espinar, B. Gschwind, Z. Qu, L. Wald, M. Schroedter-Homscheidt, C. Hoyer-Klick, A. Arola, A. Benedetti, J.W. Kaiser, J.-J. Morcrette, McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions, *Atmos. Meas. Tech.* 6 (2013) 2403–2418. URL: <https://www.atmos-meas-tech.net/6/2403/2013/>.
- [51] J. Fiorucci, T. Pellegrini, F. Louzada, F. Petropoulos, A. Koehler, Models for optimising the theta method and their relationship to state space models, *Int. J. Forecast.* 32 (2016), <https://doi.org/10.1016/j.ijforecast.2016.02.005>.
- [52] J. Castle, J. Doornik, D. Hendry, Forecasting principles from experience with forecasting competitions, *Forecasting* 3 (2021) 138–165, <https://doi.org/10.3390/forecast3010010>.
- [53] R. Hyndman, Y. Khandakar, Automatic time series forecasting: the forecast package for r, *J. Stat. Software* 26 (2008), <https://doi.org/10.18637/jss.v027.i03>.